Et aussi **sur le Web**

- **Tous les items** en version PDF
- **Tests** de lecture
- Qu'est-ce qui peut tomber à l'examen ? **Les réponses**
- **Des photothèques**
- **Des quiz**
- **Des vidéos**
- **Des interviews**

RÉFÉRENCES UNIVERSITAIRES

Les « datasets » : une triple évolution

Jean-Michel Chabot
Faculté de médecine, Marseille

En avril 2014, une interrogation Google avec les termes « dataset » ou « set minimal de données cliniques » renvoyait à près de 300 000 résultats (ce qui n'est pas considérable). Parmi ces résultats, on retrouve avec une fréquence particulière une publication de 2011, par une équipe de l'université du Minnesota, jointe aux Actes de la réunion annuelle de l'*American Medical Informatics Association* (association professionnelle, constituée voilà plus de 30 ans et dont la mission est d'exercer le leadership afin d'assurer, *via* la recherche, la formation et les bonnes pratiques en informatique, la mutation du système de santé [**http://www.amia.org/ about-amia/mission-and-history**]).

Cette publication[1] par une équipe pluridisciplinaire faisait le constat de l'importance croissante de la notion de dataset et déplorait – avec l'intention immédiate d'y remédier – l'absence d'indications et de méthode afin de développer de tels « sets de données cliniques ».

Des auteurs français, se référant volontiers à la publication de leurs collègues du Minnesota, ont proposé la définition suivante du « minimum dataset » : jeu de données commun, nécessaire pour remplir un objectif déterminé (**http://fr.slideshare.net/fullscreen/isyrare/dfinition-du-minimum-dataset-bamara-bndmr/1**). Ces mêmes auteurs, dans un article consacré à l'interopérabilité des bases de données, avancent une définition fonctionnelle dans les termes suivants : *Le concept de « minimum dataset » s'est développé pour mettre en œuvre des études à grande échelle dans le domaine de la santé avec des données comparables. Des minimums datasets sont généralement proposés pour des études spécifiques … et pour la mise en œuvre de registres* (**http://www.bndmr.fr/ wp-content/uploads/2014/02/MDS-IC2012-CHOQUET.pdf**).

Dès lors, on peut aisément préciser la finalité recherchée par l'élaboration de ces datasets : il s'agit de déterminer un nombre minimal de données – cliniques, s'agissant de santé et de maladie, à l'échelon de l'individu ou de groupe(s) de malades – qui permettent, avec un bon degré de validité, de caractériser un diagnostic, de suivre l'évolution et d'échanger entre professionnels intéressés/impliqués.

L'informatisation croissante des professionnels de santé et le regroupement de ces professionnels pour constituer des équipes plus ou moins formalisées – et les échanges d'informations que cela implique – constituent évidemment un contexte favorable pour l'élaboration de ces datasets. C'est de ce point de vue que l'équipe de l'université du Minnesota s'est placée pour proposer son cadre méthodique de production de datasets.

Selon ce cadre, un set minimal de données cliniques s'inscrit dans les 5 dimensions suivantes :

- c'est un ensemble de données standardisées cohérentes et explicitement articulées ;
- développé en utilisant une approche empirique explicite pour définir et nommer les constructions cliniques pertinentes ;
- conçu pour représenter et capturer les données de façon optimale à l'interface patient/terminal du système d'information (à portée du médecin) ;
- implémenté de telle façon qu'il peut être intégré avec d'autres sets de données (sur d'autres maladies) ;

TOUS DROITS RESERVES - LA REVUE DU PRATICIEN

– orienté vers l'acquisition de connaissances exploitables sur un terminal du système d'information.

Dans des termes plus profanes, c'est donc un ensemble de données cliniques qui permet de suivre au mieux l'évolution d'un malade. Cet ensemble de données cliniques est volontairement limité aux données sur lesquelles la plupart des analyses de l'état du malade sont établies et aux décisions qui en découlent. En pratique, un dataset pour une maladie commune comme la BPCO ou l'insuffisance rénale comporte moins de 10 données cliniques.

Établir les données qui constituent un dataset nécessite de croiser une expertise (fondée sur une approche de type EBM, c'est-à-dire associant l'analyse de la littérature et attentive au contexte d'exercice et aux expressions du patient) et des retours d'expériences comme peuvent le faire des praticiens de terrain. Et même s'il est stabilisé, un dataset a vocation à être régulièrement actualisé et amélioré.

Ce sont sur les données réunies dans un dataset que le travail en équipe au sein d'un regroupement professionnel (maison ou pôle de santé) s'élabore. Ainsi, on peut considérer qu'un dataset constitue la forme la plus concise, la plus fonctionnelle, d'un protocole partagé pluriprofessionnel.

Dès lors que les données sont limitées en nombre, il est possible de les « structurer », et donc d'organiser leur regroupement et leur combinaison (qui s'opèrent automatiquement). Ces regroupements et ces combinaisons conduisent à deux grandes fonctionnalités que les professionnels utilisent presque constamment au cours de leur exercice :

– d'abord une assistance « en temps réel » à l'exercice (par la production de mémos ou d'alertes qui peuvent être appelés à l'écran – y compris d'un smartphone – dès lors que la situation d'un malade y incite) ;
– ensuite par une analyse immédiate puis différée des parcours des malades, autant en termes individuels qu'en termes de cohorte.

De surcroît, ces datasets, les données qui les constituent et les analyses qui peuvent en être faites, sont de première utilité quand un malade doit être référé à un professionnel ou à un plateau technique de deuxième recours (en particulier par la production fiabilisée d'un « résumé de transfert ou de sortie » [***v.* encadré**])

Bref, ces datasets (dont la production a débuté dans notre monde de la santé) représentent trois évolutions en cours et constituent trois zones d'impact de la réorganisation de notre système de soins, et en particulier de soins primaires :

– d'abord dans la formation des professionnels, puisqu'au-delà des habituelles notions généralement exhaustives, exposées au cours des cycles académiques, l'identification des datasets et des éléments qui les constituent représentent une nouvelle présentation, quasiment « fonctionnelle » de l'exercice de la médecine ;
– ensuite pour l'exercice des professionnels, en particulier dans le cas des regroupements, puisque ce sont ces datasets qui constituent la base des protocoles partagés, autour desquels se construisent l'exercice en équipe et les parcours des malades ;
– enfin, sur la régulation de l'activité des professionnels, puisque les analyses et évaluations générées à partir des datasets produisent des indicateurs qui peuvent être utilisés autant pour des aspects populationnels (épidémiologie clinique ou plus simplement analyse de patientèle) que sur des aspects financiers et économiques (en particulier support à des rémunérations d'équipe ou individuelles, dans un cadre contractuel ou de performance).•

1. **Svensson-Ranallo PA, Adam TJ, Sainfort F.** A framework and standardized methodology for developing minimum clinical datasets. AMIA Summits Transl Sci Proc 2011;2011:54-8.

Extrait d'un travail de la HAS sur les synthèses de données cliniques constituant le volet médical de synthèse*

Les documents internationaux mettent en évidence trois grands types d'objectifs :

➜ une « synthèse médicale » intégrée au dossier patient et destinée au praticien (le médecin habituel, l'étudiant ou le remplaçant) lui permettant, en consultation, d'avoir en un coup d'œil une présentation de la situation clinique globale du patient ; c'est la valeur d'usage de la « synthèse » proposée par le *Royal Australian College of General Practitioners*, le *College of Physicians and Surgeons of Ontario* et de la synthèse intégrée de Stelman sur le site de l'AAFP (*American Academy of Family Physician*) ;

➜ des « *Patient Summaries* » pour alimenter les dossiers médicaux partagés (*Electronic Health Record-like System*) avec, d'une part, des versions minimales pour faire face aux situations d'urgence : c'est le cas de : *Summary Care Record* au Royaume-Uni, *Emergency Care Summary* en Écosse, epSOS '*Minimum Dataset*' entre pays européens ; et, d'autre part, des versions avec davantage de contenu pour permettre une meilleure coordination des soins dans le cadre de soins non programmés comme le SumEHR en Belgique, les *Condition Specific Summaries* qui s'ajoutent au *Minimum Dataset* en Finlande, le *National Patient Summary* suédois (NPÖ), le *Patient Summary Extended* du projet epSOS ;

➜ des synthèses médicales destinées à être transférées électroniquement entre professionnels de santé ou entre professionnels et établissements de santé. Les documents analysés décrivent le contenu métier mais aussi les règles d'interopérabilité. Sont retenus ici :

– le « *Patient Summary Record* » qui contient les données minimales à transmettre dans le cadre d'un échange *via* SI entre médecins, défini par le *Department of Health and Human Services* américain ;
– le « *Medical Summary* » qui est le contenu minimum d'une synthèse médicale défini par l'IHE PCC (*Integrating the Healthcare Entreprise Patient Care Coordination*) ;
– le « *Referral Summary* » qui définit les données que devraient contenir une synthèse accompagnant la demande d'un avis auprès d'un autre spécialiste selon l'IHE PCC.

*** http://www.has-sante.fr/portail/upload/docs/application/pdf/2013-04/synthese_medicale_quel_contenu_analyse_bibliographique.pdf**

TOUS DROITS RESERVES - LA REVUE DU PRATICIEN